HHS Public Access

Author manuscript

Mov Disord. Author manuscript; available in PMC 2024 December 01.

Published in final edited form as:

Mov Disord. 2023 December ; 38(12): 2230–2240. doi:10.1002/mds.29605.

Bidirectional transcription at the *PPP2R2B* gene locus in spinocerebellar ataxia type 12

Chengqian Zhou, M.S.[1], **Hans B. Liu, Sc.M.**[1], **Fatemeh Jahanbakhsh, Ph.D.**[2], **Leon Deng, B.A.**[1], **Bin Wu, Ph.D.**[2,5], **Mingyao Ying, Ph.D.**[3,4], **Russell L. Margolis, M.D.**[1,3], **Pan P. Li, Ph.D.**[1,*]

[1]Department of Psychiatry and Behavioral Sciences, Division of Neurobiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[2]Department of Biophysics and Biophysical Chemistry, Johns Hopkins University, Baltimore, Maryland, USA

[3]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

[4]Hugo W. Moser Research Institute at Kennedy Krieger, Baltimore, Maryland, USA

[5]Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Abstract

BACKGROUND: Spinocerebellar ataxia type 12 (SCA12) is a neurodegenerative disease caused by expansion of a CAG repeat in the *PPP2R2B gene*.

OBJECTIVE: Here we tested the hypothesis that the *PPP2R2B antisense* (*PPP2R2B-AS1*) transcript containing a CUG repeat is expressed and contributes to SCA12 pathogenesis.

METHODS: Expression of *PPP2R2B-AS1* transcript was detected in SCA12 human induced pluripotent stem cells (iPSCs), iPSC-derived NGN2 neurons, and SCA12 knock-in mouse brains using strand-specific RT-PCR (SS-RT-PCR). The tendency of expanded *PPP2R2B-AS1* (*expPPP2R2B-AS1*) RNA to form foci, a marker of toxic processes involving mutant RNAs, was examined in SCA12 cell models by fluorescence *in situ* hybridization. The apoptotic effect of *expPPP2R2B-AS1* transcripts on SK-N-MC neuroblastoma cells was evaluated by caspase 3/7 activity. Western blot was used to examine the expression of repeat associated non-ATG-initiated (RAN) translation of *expPPP2R2B-AS1* transcript in SK-N-MC cells.

*Corresponding Author Pan P. Li, Department of Psychiatry and Behavioral Sciences, Division of Neurobiology, Johns Hopkins University School of Medicine, 600 N. Wolfe St., Baltimore, MD 21287, USA, Telephone: +1 4105023760; Fax: +1 4106140013; ple5@jhmi.edu.

Author contributions

P.P.L. conceptualized the study, and C.Z. and P.P.L. designed the experiments. C.Z., H.B.L., F.J.B., L.D., B.W., M.Y., R.L.M., and P.P.L. acquired and/or analyzed the data. C.Z., R.L.M. and P.P.L. wrote the manuscript. All the authors had final approval of the submitted version.

Relevant conflict of interest/financial disclosure: Nothing to report.

Potential Conflicts of Interest

The authors declare no conflict of interest.

RESULTS: The repeat region in *PPP2R2B* gene locus is bidirectionally transcribed in SCA12 iPSCs, iPSC-derived NGN2 neurons, and SCA12 mouse brains. Transfected *expPPP2R2B-AS1* transcripts induce apoptosis in SK-N-MC cells, and the apoptotic effect may be mediated, at least in part, by the RNA secondary structure. The *expPPP2R2B-AS1* transcripts form CUG RNA foci in SK-N-MC cells. *expPPP2R2B-AS1* transcript is translated in the Alanine ORF via repeat-associated non-ATG (RAN) translation, which is diminished by single nucleotide interruptions within the CUG repeat, and MBNL1 overexpression.

INTERPRETATION: These findings suggest that *PPP2R2B-AS1* contributes to SCA12 pathogenesis, and may therefore provide a novel therapeutic target for the disease.

Introduction

Spinocerebellar ataxia type 12 (SCA12) is an autosomal dominant neurodegenerative disease, among the more common SCAs in India[1–3], characterized by tremor, gait abnormalities, and neuropsychiatric syndromes[3]. Neuropathologically, the single SCA12 brain available for study revealed both cerebral cortical and cerebellar atrophy, with a noted loss of Purkinje cells[4], in line with the cerebral cortical and cerebellar atrophy consistently observed by CT and MRI[3, 5].

SCA12 is caused by a CAG/CTG expansion mutation in exon 7 of *PPP2R2*B, a gene encoding regulatory units of protein phosphatase 2A (PP2A)[6]. Normal alleles carry 4 to 31 triplets[7], whereas disease alleles carry 43 to 78 triplets[2]. No polyGln aggregates were detected in the available formalin-fixed SCA12 brain, suggesting that polyGln may not contribute to SCA12 pathogenesis[4]. While all normal alleles and the majority of expanded alleles are uninterrupted, expanded alleles with single nucleotide interruptions within the CAG/CTG repeat have been reported in SCA12 patients of Uyghur ethnicity with much milder symptoms[8], suggesting that the hairpin RNA secondary structure formed by the CAG/CUG repeat may play a contributory role in SCA12 pathogenesis. Compared with Huntington's disease (HD) and most other CAG/CTG diseases, SCA12 has a relatively longer repeat expansion, a later disease onset, and a milder course[1].

Natural antisense transcripts (NATs) that at least partially overlap with the sense strand gene appear to contribute to the pathogenesis of a number of diseases caused by trinucleotide repeat expansions, including fragile X syndrome (FXS) and fragile X-associated tremor/ataxia syndrome (FXTAS)[9], Huntington's disease (HD)[10], Huntington's disease-like 2 (HDL2)[11, 12], SCAs 2[13], 7[14] and 8[15], and C9ORF72 FTD/ALS[16]. We therefore hypothesized that in SCA12, the repeat region at *PPP2R2B* gene locus may be bidirectionally transcribed, and the antisense product may contribute to disease pathogenesis.

Here we report that the *PPP2R2B* locus is bidirectionally transcribed in SCA12 iPSCs, SCA12 iPSC derived NGN2 neurons, and SCA12 mouse brains. The antisense transcript *PPP2R2B-AS1* with a CUG repeat expansion is neurotoxic, and therefore may contribute to SCA12 pathology. This finding suggests that *PPP2R2B-AS1* is a potential therapeutic target in SCA12.

Results

Bidirectional transcription at the PPP2R2B gene locus

Modeled on our previous work on HDL2 and SCA2[11–13], diseases in which a relatively short repeat expansion triggers a devastating phenotype, probably through a combination of loss-of-function, RNA-mediated, and protein-mediated toxicity, we speculated that SCA12 may similarly involve complex mechanisms of pathogenesis. We therefore tested the possibility that an antisense transcript with an expanded CUG repeat is expressed from the opposite strand of *PPP2R2B*. In conjunction with Cedars Sinai Stem Cell Core, we generated and characterized 8 iPSC lines derived from the skin fibroblasts of three different human SCA12 patients (Table S1) using a non-integrating protocol, while three control iPSC lines generated using the same protocol as the SCA12 lines were included for this study. All the lines used in this study contained pure CAG triplets for both normal and expanded alleles. We differentiated control and SCA12 iPSCs into cortical excitatory neurons using an NGN2 overexpression protocol[17, 18]. Using total RNA extracted from control and SCA12 iPSCs as well as iPSC-derived NGN2 neurons, we performed strand-specific reverse transcription PCR (SS-RT-PCR, Fig. S1)[13] with linkered (LK) primers flanking the repeat in *PPP2R2B* exon 7 (F1 and R1, Fig. 1A). We detected the expression of *PPP2R2B* antisense transcripts from both normal and expanded alleles from SCA12 iPSCs or iPSC derived NGN2 neurons, and from the normal alleles in the corresponding controls (Fig. 1B and 1C). In keeping with HUGO guidelines, we named the gene expressing this transcript *PPP2R2B* antisense 1, *PPP2R2B-AS1*. The identity of the transcript was confirmed by sequencing (Table S2). Next, we examined the expression of *PPP2R2B-AS1* in a novel humanized SCA12 knock-in mouse model that was generated by replacing the mouse *PPP2R2B* exon 2 containing a short, interrupted repeat with its human counterpart, exon 7, with either 10 or 80 CAG triplets (KI-10 or KI-80; Li and Margolis, unpublished results). SS-RT-PCR detected expanded *PPP2R2B-AS1* (*expPPP2R2B-AS1*) transcripts in the cortices and cerebella of three-month-old male homozygous SCA12 KI-80 mice, while as expected, normal *PPP2R2B-AS1* (*nPPP2R2B-AS1*) transcripts were expressed in the age- and sex- matched homozygous SCA12 KI-10 or wildtype (WT) cortices or cerebella (Fig. 1D). The *PPP2R2B-AS1* transcript fragment detected by SS-RT-PCR contains 44bp upstream and 19bp downstream of the CUG repeat, and includes no ATG start codons.

3’ RACE was performed to characterize the 3’ end of the *PPP2R2B-AS1* transcript in control human brain, and we detected additional 3’ sequence (145bp downstream of the CTG repeat; without ATGs) followed by a polyA tail (Table S2), indicating that the *PPP2R2B-AS1* transcript is indeed polyadenylated (Fig. 1A). Despite multiple attempts at 5’ RACE, the 5’ end remains uncharacterized, presumably because high GC content prevented successful PCR amplification. Nevertheless, three SP1 binding sites are predicted in the region 98~324 bp upstream of the CTG repeat[19], indicating potential transcription initiation sites for the *PPP2R2B-AS1* transcript. Future full-length RNAseq using control human brain tissue, or SCA12 KI mouse brain, may reveal the entire sequence and possible splicing isoforms of *PPP2R2B-AS1*, and guide experiments to confirm the presence and properties of a functional promoter for *PPP2R2B-AS1*, and suggest likely ORFs (if any) within the full length *PPP2R2B-AS1* transcript.

expPPP2R2B-AS1 induces apoptosis in neuroblastoma SK-N-MC cells

Transcripts containing expanded CUG repeats contribute to toxicity in DM1, HDL2, and SCA2[11, 13], and perhaps other CAG/CTG diseases[20]. We therefore examined whether *expPPP2R2B-AS1* with a CUG repeat expansion within the physiological range of adult-onset SCA12 induces apoptosis in neuronal-like cells. *PPP2R2B-AS1-(CTG)n* constructs expressing *PPP2R2B-AS1* transcript fragments with either 10 or 73 CUG triplets and flanking regions extending 44 bp upstream and 145 bp downstream from the repeat were transfected into SK-N-MC neuroblastoma cells. Caspase 3/7 activity was assayed 72 hours after overexpression of the exogenous *PPP2R2B-AS1* in SK-N-MC cells, and was compared to that of the exogenous *ATXN2 antisense* (*ATXN2-AS)* transcript fragment with 44 (vs. 22) CUG triplets, previously shown to be associated with SCA2 pathogenesis and toxic to SK-N-MC cells[13]. *PPP2R2B-AS1-(CTG)73* triggered significant Caspase 3/7 activation that was similar to that of the *ATXN2-AS-(CTG)44* transcript and was about 1.5 times more than *PPP2R2B-AS1-(CTG)10 and ATXN2-AS-(CTG)22* (Fig. 2A).

Individuals from a single SCA12 pedigree of Uyghur ethnicity are predicted to have an expanded SCA12 allele that would express *expPPP2R2B-AS1* with 1–3 single nucleotide interruptions in the repeat[8]; these individuals have a mild form of SCA12[8], suggesting that interruptions in the CUG repeat of *expPPP2R2B-AS*1 may reduce neurotoxicity. It has been suggested that CUG transcript toxicity is dependent on the hairpin structure formed by the repeat[13], the stem of which sequesters RNA-binding proteins (RBPs), i.e. MBNL1; we therefore hypothesized that the *expPPP2R2B-AS1* transcripts with an interrupted repeat, as found in the Uyghur family, may induce apoptosis less than *expPPP2R2B-AS1* transcripts with an uninterrupted CUG repeat of the same length. To test this hypothesis, we introduced three single nucleotide mutations into the repeat of *PPP2R2R-AS1-(CTG)73*, in order to obtain *PPP2R2B-AS1-Int(CTG)73*, with a repeat composition of $CTG)_{23}CTC(CTG)_{14}CCG(CTG)_{13}TTG(CTG)_{20}$, similar to the interrupted *expPPP2R2B-AS1* transcript predicted in the Uyghur population. Inserting interruptions indeed decreased the apoptosis triggered by the *expPPP2R2B-AS1* in SK-N-MC cells, as measured by Caspase 3/7 activity assay (Fig. 2B). The levels of overexpressed transcripts in SK-N-MC cells were measured by quantitative RT-PCR (qRT-PCR; Fig. 2C), indicating that though the same amount (1ug) of dsDNA plasmid of each construct was transfected, the *PPP2R2B-AS1-(CUG)73* transcript was expressed at a much lower level than *PPP2R2B-AS1-Int(CUG)73* and *PPP2R2B-AS1-(CUG)10*, suggesting that the *expPPP2R2B-AS1*'s effect in inducing apoptosis, observed in the Caspase 3/7 assay, may be underestimated. We speculate that the low level of *PPP2R2B-AS1-(CUG)73* expression may be the consequence of nuclear retention, perhaps in foci (see below). This set of experiments suggests that the ability of *expPPP2R2B-AS1* to induce apoptosis is dependent on both repeat length and repeat composition, potentially related to the secondary structure formed by the repeat structures.

expPPP2R2B-AS1 transcripts form CUG RNA foci

Repeat-containing mutant transcripts form RNA foci in all CUG/CAG diseases in which RNA neurotoxicity contributes to pathogenesis [21–23]. RNA foci are also a hallmark of RNA toxicity in neurodegenerative repeat diseases associated with other types of repeats,

including the ATTCT pentamer expansion that causes spinocerebellar ataxia type 10 (SCA10) [24] and the GGGGCC hexamer expansion in c9orf72 that is associated with ALS [25].

We therefore sought to detect similar foci in *PPP2R2B-AS1* overexpressing SK-N-MC cells by fluorescence *in situ* hybridization (FISH) using a 20-mer CAG LNA/DNA probe that binds to the CUG repeat[26]. CUG RNA foci were absent in SK-N-MC neuroblastoma cells that overexpress *PPP2R2B-AS1-(CTG)10* (Fig. 3A), but were not uncommon in cells overexpressing *PPP2R2B-AS1-(CTG)73* (Fig. 3B) or *PPP2R2B-AS1-Int(CTG)73* (Fig. 3C), as quantified in Fig. 3M. *PPP2R2B-AS1-(CUG)73* and *PPP2R2B-AS1-Int(CUG)73* formed similar number of foci, except that the % cells with 10+ foci was higher in *PPP2R2B-AS1-(CUG)73* cells, compared with *PPP2R2B-AS1-Int(CUG)73* cells. This set of experiments demonstrates that *expPPP2R2B-AS1* transcripts form RNA foci.

Consistent with previous reports that MBNL1, a nuclear splicing factor that aberrantly interacts with CUG repeat transcripts, colocalizes to CUG RNA foci[13, 27, 28], we also found that although exogenously expressed GFP-MBNL1 is diffusely localized to the nuclei of SK-N-MC cells co-transfected with *PPP2R2B-AS1-(CTG)10* (Fig. 3F), co-expression of GFP-MBNL1 and *PPP2R2B-AS1-(CTG)73 or PPP2R2B-AS1-Int(CTG)73* results in co-localization of MBNL1 with CUG RNA foci (Fig. 3I and 3L). While the toxicity of RNA foci themselves remains to be determined, the colocalization of RNA foci with MBNL1 in this model is suggestive of a pathogenic role of *expPPP2R2B-AS1* in SCA12.

RAN translation of the expPPP2R2B-AS1 transcript

The experimentally-defined transcript sequence of *PPP2R2R-AS1* (the sequence used in our experiments) does not include ATG start codons upstream of the CUG repeat and would not be predicted to yield protein products via canonical ATG-initiated translation. However, hairpin-forming expanded CUG repeats can also be translated in the absence of ATG start codon through repeat-associated non-ATG translation (RAN translation) [29]. To determine if *PPP2R2B-AS1* has the potential of undergoing RAN translation, and if RAN-translated protein fragments with expanded amino acid tracts lead to neurotoxicity, we cloned an *PPP2R2B-AS1* fragment containing multiple upstream stop codons, and a CTG repeat expansion with flanking regions extending 44 bp upstream and 26 bp downstream from the repeat into a vector with tags for each of the three open reading frames (Fig. 4A), to obtain the *PPP2R2B-AS1-(CTG)n-3T* and *PPP2R2B-AS1-Int(CTG)73–3T* constructs. Immunostaining and western blot using antibodies against each epitope tag detected a polyalanine-containing RAN product from SK-N-MC cells that overexpress *PPP2R2B-AS1-(CTG)73–3T* (Fig. 4C and 4N) or *PPP2R2B-AS1-Int(CTG)73–3T* (Fig. 4D and 4N), but not polycystine- (Fig. 4F–H) or polyleucine-containing (Fig. 4J–L) RAN products. Quantification of polyalanine RAN protein levels reveals that the *PPP2R2B-AS-Int(CTG)73–3T* construct produced less polyalanine RAN protein than the *PPP2R2B-AS1-(CTG)73–3T* construct (Fig. 4O), indicating that the expression level of polyalanine RAN product is likely dependent on the secondary structure of the *expPPP2R2B-AS1* transcript.

MBNL1 decreases RAN translation of the *expPPP2R2B-AS1* transcript.

Given our findings that GFP-MBNL1 is sequestered to the CUG RNA foci formed by the *expPPP2R2B-AS1* transcripts (Fig. 3G–3L), and previous reports that overexpression of MBNL1 decreased protein translation of transcripts containing expanded *CAG/CUG* (*expCAG/CUG)* repeat[30], we therefore hypothesized that the MBNL1, via its interaction with the expanded *CUG* repeat, may inhibit the translation of polyalanine RAN protein from the *expPPP2R2B-AS1* transcript. In support of our hypothesis, we found that overexpression of GFP-MBNL1 (vs. GFP alone) decreased the expression of polyalanine RAN protein from the *PPP2R2B-AS1-(CUG)73–3T* transcript in SK-N-MC cells (Fig. 4P and Q). The MBNL1 protein consists of four zinc finger (ZNF) domains, an unstructured domain, and a C-terminal splicing domain containing the nuclear localization signal (NLS)[31, 32] (Fig. 5A). It was previously reported that the CUG repeats interact with MBNL1 via its N-terminal ZNF domains. To ascertain whether the effect of MBNL1 on polyalanine RAN protein production is dependent on the ZNF domains, we obtained constructs encoding GFP-tagged truncated MBNL1 proteins lacking either the C-terminal splicing domain (GFP-MBNL1-Δ251–388), or all four ZNF domains in the N-terminus (GFP-MBNL1-Δ1–240) (Fig. 5A). GFP-MBNL1 (Fig. 5B and 5E) and GFP-MBNL1-Δ1–240 (Fig. 5J and 5M) are both diffusedly localized in the nucleus, while GFP-MBNL1-Δ251–388, due to the lack of C-terminal domain containing the NLS, is expressed diffusely in nucleus and cytoplasm (Fig. 5F and 5I). GFP tagged full-length or truncated MBNL1 and *PPP2R2B-AS1* were overexpressed in SK-N-MC cells, and FISH experiments revealed that similar to full length GFP-MBNL1 (Fig. 5C–E), GFP-MBNL1-Δ251–388 (truncated MBNL1 containing ZNFs) is still colocalized with CUG RNA foci formed by *PPP2R2B-AS1-(CUG)73* transcripts (Fig. 5G–I), while GFP-MBNL1-Δ1–240 was detected diffusely in the nucleus with no colocalization with CUG RNA foci (Fig. 5K–M), confirming that the ZNF domains of MBNL1 are responsible for interacting with transcripts containing expanded CUG repeats. Next, we examined the effect of truncated MBNL1 on the production of polyalanine RAN protein from the *expPPP2R2B-AS1–3T* transcripts. While overexpression of GFP-MBNL1-Δ251–388 decreased the production of the polyalanine RAN protein from *PPP2R2B-AS1-(CUG)73–3T* transcripts, mimicking the effect of full length GFP-MBNL1, overexpression of GFP-MBNL1-Δ1–240 (lacking all the four ZNF domains) did not affect polyalanine RAN protein production (Fig. 5N and 5O), indicating that ZNF domains are necessary for MBNL1 to inhibit RAN translation of the *expPPP2R2B-AS1* transcript. However, not all four ZNF domains may be necessary, as overexpression of GFP-MBNL1-Δ12–46, a GFP tagged MBNL1 lacking the first ZNF only (Fig. S2A), also decreased the RAN polyalanine production from the *PPP2R2B-AS1-(CUG)73–3T* transcripts (Fig. S2B and S2C), mimicking the effect of full length GFP-MBNL1. To ascertain whether the inhibitory effect of MBNL1 ZNF domains on translation of the *expanded CUG repeat* transcripts also applies to ATG-initiated translation, we included a construct of *JPH3* exon 2A containing an expanded CTG repeat (relevant to HDL2), *JPH3 Ex2A-(CTG)55*, which encodes a protein containing a long polyalanine tract via ATG translation[11, 30]. Similar to the RAN translation of the *PPP2R2B-AS1-(CUG)73–3T* transcripts, overexpression of GFP-MBNL1 (Fig. 5P), GFP-MBNL1-Δ251–388 (Fig. 5P), GFP-MBNL1-Δ12–46 (Fig. S2B), but not that of GFP-MBNL1-Δ1–240 (Fig. 5P) decreased the expression of polyalanine protein from the

JPH3 Ex2A-(CTG)55 (Fig. 5Q and Fig. S2C), indicating that MBNL1 ZNF domains have inhibitory effects on the ATG-translation of the expanded *CUG* repeat as well.

Discussion

We demonstrate that both normal and expanded CUG repeat-containing *PPP2R2B-AS1* transcripts are expressed in human SCA12 iPSCs and iPSC derived cortical neurons, as well as in a novel SCA12 knock-in mouse model. The *expPPP2R2B-AS1* transcripts form foci that sequester the splicing factor MBNL1. Further, we show that the *expPPP2R2B-AS1* with a CUG repeat size within the physiological range for adult-onset SCA12 is toxic to neuronal-like SK-N-MC neuroblastoma cells. Finally, the *expPPP2R2B-AS1* transcript expresses polyalanine RAN protein which is decreased by single nucleotide interruptions within the CUG repeat, overexpression of full-length MBNL1, or overexpression of a truncated MBNL1 containing ZNF domains. We conclude that the *expPPP2R2B-AS1* transcripts may play a role in the pathogenesis of SCA12.

In DM1, and HD, normal, but not expanded antisense transcripts are expressed at the respective disease loci in human tissue[10, 11, 33]. In HDL2, expanded antisense transcript was detected in a BAC-HDL2 transgenic mouse model[12], but not in frozen human postmortem brains[11]. In SCA2 and SCA8, the antisense transcript with the expanded repeat is expressed at significant levels[13, 15]. Only one human SCA12 postmortem brain (formalin fixed) was available for this study[34], and SS-RT-PCR detected only the antisense allele with the normal repeat (Fig. S3). We speculate that detection of the low-abundant antisense transcript with the expanded repeat likely requires better quality RNA than is available from the formalin fixed brain; in the future, we will detect *PPP2R2B-AS1* transcript in frozen (not formalin fixed) human SCA12 postmortem brains, as they become available. Nevertheless, *expPPP2R2B-AS1* was detected in undifferentiated and differentiated SCA12 iPSCs (Fig. 1B and 1C), and in the cortices and cerebella of SCA12 knock-in mice (Fig. 1D), indicating the potential relevance of *expPPP2R2B-AS1* to SCA12.

NATs frequently regulate sense strand transcript expression[37,23], as in the case of *HTT-AS* regulation of *HTT*[10, 35]. Our preliminary data suggests that overexpression of a *PPP2R2B-AS1* fragment *in trans* does not have a substantial effect on the expression of endogenous *PPP2R2B* transcript in SK-N-MC neuroblastoma cells (Fig. S4). Whether *PPP2R2B-AS1* affects the expression of *PPP2R2B in cis* remains to be determined.

Overexpressed *expPPP2R2B-AS1* fragments readily aggregate into RNA foci in SK-N-MC cells (Fig. 3). The role of these foci in the pathogenesis of other repeat expansion diseases remains controversial[13, 36, 37]; it is not fully established whether the foci themselves are protective, an epiphenomenon or neurotoxic. Similarly, the direct effect of the foci in SCA12 pathogenesis awaits further elucidation.

Our data nevertheless indicate a toxic role of *expPPP2R2B-AS1* transcripts. Although the mechanism of *expPPP2R2B-AS1* toxicity is not entirely clear, we speculate that the toxicity may be, in part, mediated by “pure” RNA toxicity, potentially *via* the aberrant interaction between the expanded CUG repeat and RNA binding proteins (RBPs), and in support of this

speculation, we observed colocalization of *expPPP2R2B-AS1* foci and MBNL1, an RBP that is involved in RNA splicing. Future *in vitro* biotinylated RNA pull down experiments[38] may reveal additional RBPs that preferentially interact with the expanded CUG repeat on *expPPP2R2B-AS1* transcripts.

In addition, our data leads us to speculate that RAN translation of expanded polyalanine tracts from the *expPPP2R2B-AS1* transcript contributes to SCA12 neurotoxicity, as proteins or peptides containing long polyalanine tract have been previously shown to be toxic in *in vitro* cell culture or *drosophila* models[39–41]. In the future, we will detect the expression of RAN polyalanine products in the SCA12 knock-in mouse brains, when the specific antibodies become available. The interrupted *PPP2R2B-AS1-Int(CUG)73* transcript is less toxic and also produces less RAN polyalanine protein, indirect evidence that RAN translation of expanded polyalanine tracts contributes to the transcript toxicity. Since all four alanine-encoding codons (GCT, GCC, GCA, and GCG) are GC-rich, it is not possible to design a long polyalanine-encoding transcript sequence without a GC-rich hairpin structure[40], making it challenging to evaluate the relative contributions of the "pure" RNA toxicity and RAN polyalanine protein mediated toxicity. Future experiments employing additional tools, such as auxin-inducible degrons[42–44], to specifically degrade RAN polyalanine protein, while leaving behind the *expPPP2R2B-AS1* transcript, may help answer this question.

Given that somatic instability has been implicated in microsatellite repeat expansion disorders such as HD[45, 46], it is possible that somatic expansion of the pure CAG repeat may exist in SCA12 brains, and that the single nucleotide interruptions within CAG repeat, associated with the Uyghur SCA12 with mild diseases, may stabilize the repeat, leading to milder disease phenotypes compared with SCA12 patients with expanded pure CAG repeats. Future experiments using frozen human SCA12 postmortem brains may help clarify the role of somatic instability in SCA12 pathogenesis.

Our qRT-PCR results indicate that exogenous *expPPP2R2B-AS1* transcripts are expressed at a lower level than *PPP2R2B-AS1* transcripts with a normal repeat, potentially due to that transcripts sequestered within CUG RNA foci may not be readily extractable, or that the presence of the expanded CUG repeat (i.e., hairpin structure) interferes with transcript expression, leading to a lower transcript level.

The currently known *PPP2R2B-AS1* transcript sequence entirely overlaps with the sense *PPP2R2B* transcript (exon 7), rendering it not feasible to design antisense specific qPCR primers to measure the relative amount of antisense vs. sense transcript at the *PPP2R2B* gene locus. Future experiments identifying the full-length sequence of *PPP2R2B-AS1* will then help evaluate the contribution of the *expPPP2R2B-AS1* transcript, relative to the expanded sense *PPP2R2B* (*expPPP2R2B*) transcript to SCA12 pathogenesis. Due to the fact that antisense genes are usually expressed at a lower level than the corresponding sense genes[10, 13], we speculate that the *expPPP2R2B-AS1* may contribute relatively less to the SCA12 pathogenesis, compared with the *expPPP2R2B* transcript. Nevertheless, our results provide further evidence that the pathogenesis of CAG/CTG repeat diseases is likely "multifactorial", as bidirectionally expressed transcripts at the disease loci may have additive

or interacting effects on pathogenesis, in addition to the potential effect of the repeat on transcription or translation. This is of relevance to the development of therapy, as it suggests that targeting sense proteins/transcripts may not always be sufficient to achieve significant therapeutic benefits. Hence, suppressing the expression of *expPPP2R2B-AS1* may offer an additional or an alternative therapeutic approach to SCA12. Given our finding that overexpression of the ZNF domains of MBNL1 decreased the expression of polyalanine RAN proteins from the *expPPP2R2B-AS1* transcripts, CUG-repeat blocking peptides based on MBNL1 ZNF domains may have a neuroprotective effect. Since SCA12 belongs to a large group of diseases that are caused by a CAG/CTG repeat expansion, the data from this study may apply to other CAG/CTG repeat diseases involving additive or synergistic mechanisms of protein and RNA neurotoxicity.

Materials and Methods

A description of the materials and methods is provided in the Supporting Data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Laura P.L. Ranum, and Dr. Charles A. Thornton for the kind gift of the A8(*KKQ$_{EXP}$)-3Tf1, and GFP-MBNL1 constructs, respectively. We thank Dr. Olga Pletinikova for providing a formalin-fixed brain sample of one SCA12 patient.

Funding agencies:

This work was supported by the National Institutes of Health grants NS112687 (to P.P.L.), NS122756 (to P.P.L.), NS112796 (to P.P.L.), NS093287 (to R.L.M.), NS113820 (to B.W.), and GM136897 (to B.W.), the ABCD Charitable Trust (to R.L.M.), and the National Ataxia Foundation (to P.P.L.).

Full Financial Disclosures for the Previous 12 Months

P.P.L. receives grant funding from the National Institutes of Health, the ABCD Charitable Trust, and the Maryland Stem Cell Research Fund. R.L.M. receives grant funding from the National Institutes of Health and the ABCD Charitable Trust, and the Abramson Fund. B.W. receives grant funding from the National Institutes of Health. M.Y. receives grant funding from the National Institutes of Health and Maryland Stem Cell Research Fund. C.Z., H.B.L., F.J, and L.D. have nothing to disclose.

References

1. Choudhury S et al. Clinical Characterization of Genetically Diagnosed Cases of Spinocerebellar Ataxia Type 12 from India. Mov. Disord. Clin. Pract. 5, 39–46 (2017). [PubMed: 30363072]
2. Srivastava AK, Takkar A, Garg A & Faruq M Clinical behaviour of spinocerebellar ataxia type 12 and intermediate length abnormal CAG repeats in PPP2R2B. Brain 140, 27–36 (2017). [PubMed: 27864267]
3. Cohen RL & Margolis RL Spinocerebellar ataxia type 12: clues to pathogenesis. Curr. Opin. Neurol. 29, 735–742 (2016). [PubMed: 27748686]
4. O'Hearn EE et al. Neuropathology and Cellular Pathogenesis of Spinocerebellar Ataxia Type 12. Mov. Disord. 30, 1813–1824 (2015). [PubMed: 26340331]
5. Li H, Ma J & Zhang X Diffusion tensor imaging of spinocerebellar ataxia type 12. Med. Sci. Monit. 20, 1783–1791 (2014). [PubMed: 25274186]

6. Holmes SE et al. Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. Nat. Genet. 23, 391–392 (1999). [PubMed: 10581021]
7. O'Hearn E, Holmes SE & Margolis RL Spinocerebellar ataxia type 12. Handb. Clin. Neurol. 103, 535–547 (2012). [PubMed: 21827912]
8. LI H, LEI J, MA J, YU J & ZHANG X Gene mutation and clinical characteristics of a Chinese Uygur family with spinocerebellar ataxia type 12. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 28, 137–141 (2011). [PubMed: 21462122]
9. Ladd PD et al. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. Hum. Mol. Genet. 16, 3174–3187 (2007). [PubMed: 17921506]
10. Chung DW, Rudnicki DD, Yu L & Margolis RL A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. Hum. Mol. Genet. 20, 3467–3477 (2011). [PubMed: 21672921]
11. Seixas AI et al. Loss of junctophilin-3 contributes to Huntington disease-like 2 pathogenesis. Ann. Neurol. 71, 245–257 (2012). [PubMed: 22367996]
12. Wilburn B et al. An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron 70, 427–440 (2011). [PubMed: 21555070]
13. Li PP et al. ATXN2-AS, a gene antisense to ATXN2, is associated with spinocerebellar ataxia type 2 and amyotrophic lateral sclerosis. Ann. Neurol. 80, 600–615 (2016). [PubMed: 27531668]
14. Sopher BL et al. CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. Neuron 70, 1071–1084 (2011). [PubMed: 21689595]
15. Moseley ML et al. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat. Genet. 38, 758–769 (2006). [PubMed: 16804541]
16. Parameswaran J et al. Antisense, but not sense, repeat expanded RNAs activate PKR/eIF2alpha-dependent ISR in C9ORF72 FTD/ALS. Elife 12, 10.7554/eLife.85902 (2023).
17. Zhang Y et al. Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron 78, 785–798 (2013). [PubMed: 23764284]
18. Xue Y et al. Synthetic mRNAs Drive Highly Efficient iPS Cell Differentiation to Dopaminergic Neurons. Stem Cells Transl. Med. 8, 112–123 (2019). [PubMed: 30387318]
19. Messeguer X et al. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 18, 333–334 (2002). [PubMed: 11847087]
20. Schwartz JL, Jones KL & Yeo GW Repeat RNA expansion disorders of the nervous system: post-transcriptional mechanisms and therapeutic strategies. Crit. Rev. Biochem. Mol. Biol. 56, 31–53 (2021). [PubMed: 33172304]
21. Rudnicki DD, Margolis RL, Pearson CE & Krzyzosiak WJ Diced triplets expose neurons to RISC. PLoS Genet. 8, e1002545 (2012). [PubMed: 22383898]
22. Wojciechowska M & Krzyzosiak WJ CAG repeat RNA as an auxiliary toxic agent in polyglutamine disorders. RNA Biol. 8, 565–571 (2011). [PubMed: 21593608]
23. Urbanek MO & Krzyzosiak WJ RNA FISH for detecting expanded repeats in human diseases. Methods 98, 115–123 (2016). [PubMed: 26615955]
24. White M et al. Transgenic mice with SCA10 pentanucleotide repeats show motor phenotype and susceptibility to seizure: a toxic RNA gain-of-function model. J. Neurosci. Res. 90, 706–714 (2012). [PubMed: 22065565]
25. DeJesus-Hernandez M et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245–256 (2011). [PubMed: 21944778]
26. Mykowska A, Sobczak K, Wojciechowska M, Kozlowski P & Krzyzosiak WJ CAG repeats mimic CUG repeats in the misregulation of alternative splicing. Nucleic Acids Res. 39, 8938–8951 (2011). [PubMed: 21795378]
27. Rudnicki DD et al. Huntington's disease--like 2 is associated with CUG repeat-containing RNA foci. Ann. Neurol. 61, 272–282 (2007). [PubMed: 17387722]

28. Jiang H, Mankodi A, Swanson MS, Moxley RT & Thornton CA Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. Hum. Mol. Genet. 13, 3079–3088 (2004). [PubMed: 15496431]
29. Zu T et al. Non-ATG-initiated translation directed by microsatellite expansions. Proc. Natl. Acad. Sci. U. S. A. 108, 260–265 (2011). [PubMed: 21173221]
30. Sun X et al. Nuclear retention of full-length HTT RNA is mediated by splicing factors MBNL1 and U2AF65. Sci. Rep. 5, 12521 (2015). [PubMed: 26218986]
31. Yuan Y et al. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. Nucleic Acids Res. 35, 5474–5486 (2007). [PubMed: 17702765]
32. Konieczny P, Stepniak-Konieczna E & Sobczak K MBNL proteins and their target RNAs, interaction and splicing regulation. Nucleic Acids Res. 42, 10873–10887 (2014). [PubMed: 25183524]
33. Cho DH et al. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol. Cell 20, 483–489 (2005). [PubMed: 16285929]
34. O'Hearn EE et al. Neuropathology and Cellular Pathogenesis of Spinocerebellar Ataxia Type 12. Mov. Disord. 30, 1813–1824 (2015). [PubMed: 26340331]
35. Khaled HG et al. A high-throughput screening to identify small molecules that suppress huntingtin promoter activity or activate huntingtin-antisense promoter activity. Sci. Rep. 11, 6157–2 (2021). [PubMed: 33731741]
36. Houseley JM et al. Myotonic dystrophy associated expanded CUG repeat muscleblind positive ribonuclear foci are not toxic to Drosophila. Hum. Mol. Genet. 14, 873–883 (2005). [PubMed: 15703191]
37. Sicot G & Gomes-Pereira M RNA toxicity in human disease and animal models: from the uncovering of a new mechanism to the development of promising therapies. Biochim. Biophys. Acta 1832, 1390–1409 (2013). [PubMed: 23500957]
38. Li PP et al. RNA Toxicity and Perturbation of rRNA Processing in Spinocerebellar Ataxia Type 2. Mov. Disord. 36, 2519–2529 (2021). [PubMed: 34390268]
39. van Eyk CL, McLeod CJ, O'Keefe LV & Richards RI Comparative toxicity of polyglutamine, polyalanine and polyleucine tracts in Drosophila models of expanded repeat disease. Hum. Mol. Genet. 21, 536–547 (2012). [PubMed: 22021427]
40. Banez-Coronel M et al. RAN Translation in Huntington Disease. Neuron 88, 667–677 (2015). [PubMed: 26590344]
41. Ayhan F et al. SCA8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF3F. EMBO J. 37, e99023. doi: 10.15252/embj.201899023. Epub 2018 Sep 11 (2018). [PubMed: 30206144]
42. Wu B, Eliscovich C, Yoon YJ & Singer RH Translation dynamics of single mRNAs in live cells and neurons. Science 352, 1430–1435 (2016). [PubMed: 27313041]
43. Yesbolatova A et al. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. Nat. Commun. 11, 5701–z (2020). [PubMed: 33177522]
44. Macdonald L et al. Rapid and specific degradation of endogenous proteins in mouse models using auxin-inducible degrons. Elife 11, 10.7554/eLife.77987 (2022).
45. Kacher R et al. Propensity for somatic expansion increases over the course of life in Huntington disease. Elife 10, 10.7554/eLife.64674 (2021).
46. Monckton DG The Contribution of Somatic Expansion of the CAG Repeat to Symptomatic Development in Huntington's Disease: A Historical Perspective. J. Huntingtons Dis. 10, 7–33 (2021). [PubMed: 33579863]

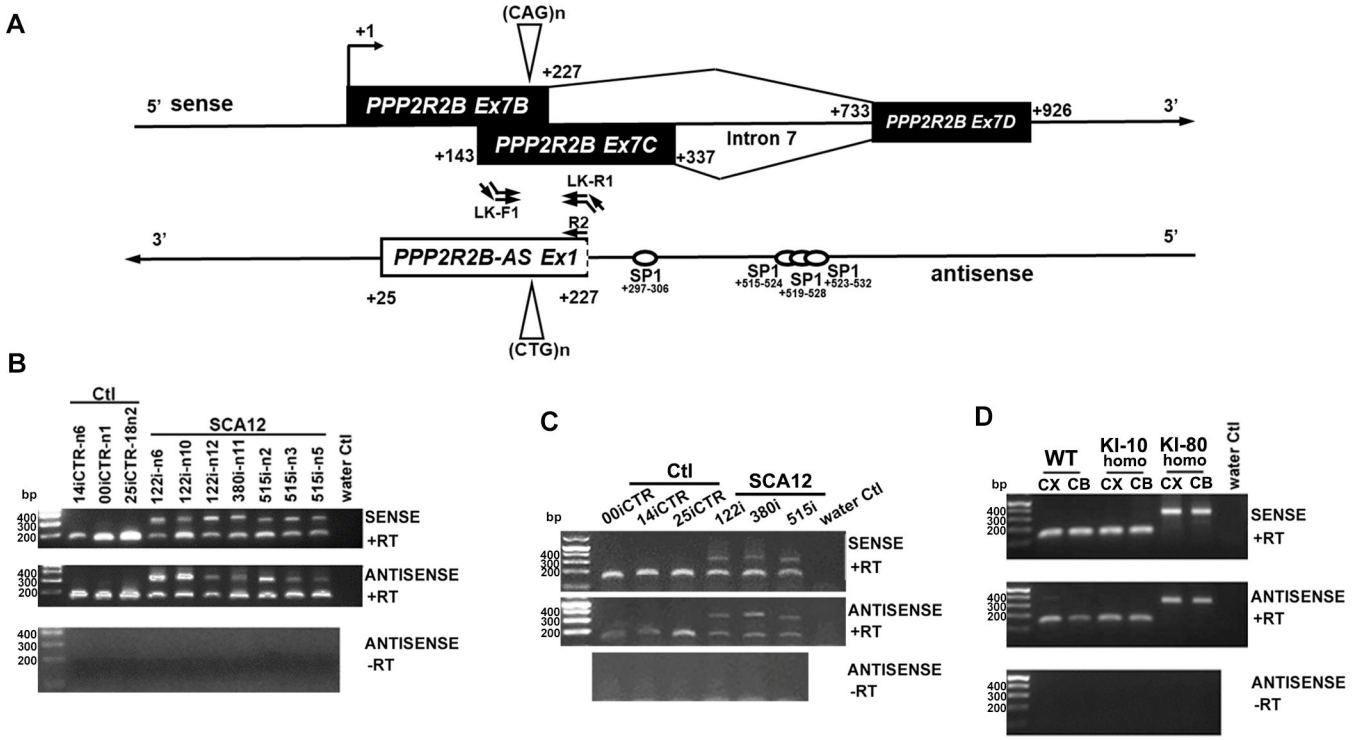

Fig. 1.

Bidirectional transcription at *PPP2R2B* gene locus. (A) *PPP2R2B-AS1* map. Primer locations for SS-RT-PCR (Linkered F1 and R1) in B-D are shown. On the sense strand, *7B7D* and *7C7D* are the two splice variants of *PPP2R2B* exon 7 that contains the CAG repeat. On the antisense strand, location of SP1 binding sites were predicted using PROMO[19]. 3' end of *PPP2R2B-AS1* was identified by 3' RACE using gene-specific primer R2 and Generacer 3' primer. The 5' end of *PPP2R2B-AS1* exon 1 is not characterized. (B-D) Detection of normal and expanded *PPP2R2B* and *PPP2R2B-AS1* by SS-RT-PCR in human iPSCs (B), iPSC derived cortical excitatory neurons differentiated by NGN2 overexpression (C), and cortices (CX) and cerebella (CB) of a humanized SCA12 knock-in mouse model with either 10 or 80 repeats (KI-10 or KI-80). PCR bands amplified from two normal alleles in human control iPSCs or iPSC derived neurons could not be separated on an agarose gel due to size similarity. Only homozygous (homo) mice were included. N=3 independent experiments; representative gel images are shown. PCRs using water (no template) as negative controls. +RT indicates RT-PCR products, while no reverse transcription (−RT) control is a mock reverse transcription containing all the RT-PCR reagents, except the reverse transcriptase.

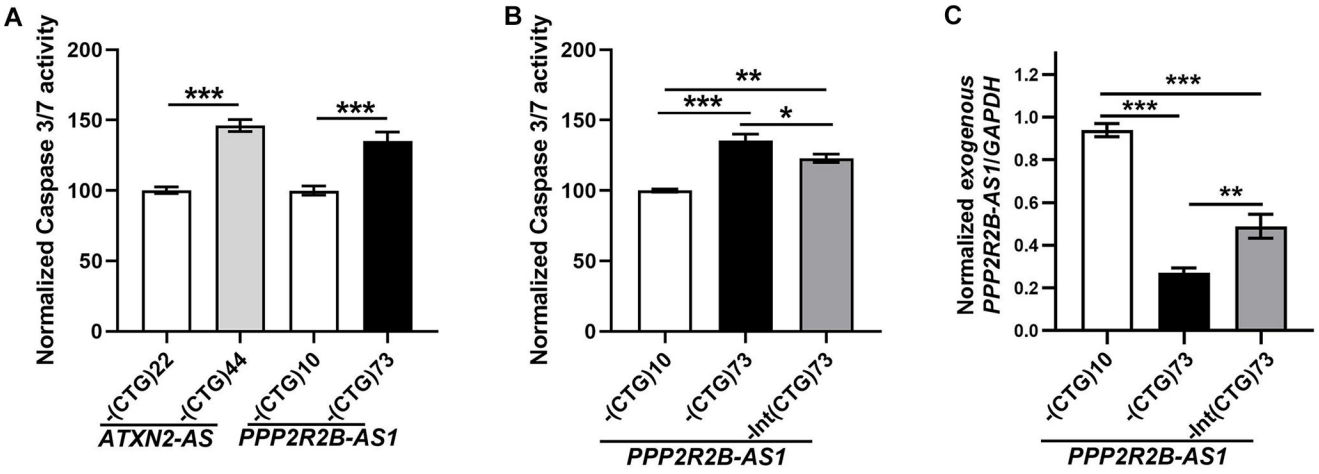

Fig. 2.

Overexpression of an expanded *PPP2R2B-AS1* transcript is toxic to SK-N-MC cells. (A) Caspase 3/7 assay indicates toxicity of an expanded *PPP2R2B-AS1-(CUG)73* transcript compared to a normal *PPP2R2B-AS1-(CUG)10* transcript. *ATXN2-AS* constructs are included as positive controls. N=5 biological replicates. The Caspase 3/7 activity in *ATXN2-AS-(CTG)22* or *PPP2R2B-AS1-(CTG)10* transfected SK-N-MC cells was normalized to 100. (B) Single nucleotide interruptions in the CUG repeat decrease the toxicity of expanded *PPP2R2B-AS1* transcript. N=4 biological replicates. The Caspase 3/7 activity in *PPP2R2B-AS1-(CTG)10* transfected SK-N-MC cells was normalized to 100. (C) Relative expression level of exogenous *PPP2R2B-AS1* transcripts in SK-N-MC cells normalized by *GADPH* level. qPCR primers used to measure exogenous *PPP2R2B-AS1* are located on the vector backbone. N=3 biological replicates. The *PPP2R2B-AS1*/*GAPDH* in *PPP2R2B-AS1-(CTG)10* transfected SK-N-MC cells was normalized to 1. Mean ± SEM are shown. One-way ANOVA and Tukey posthoc test, *p<0.05, **p<0.01, ***p<0.001.

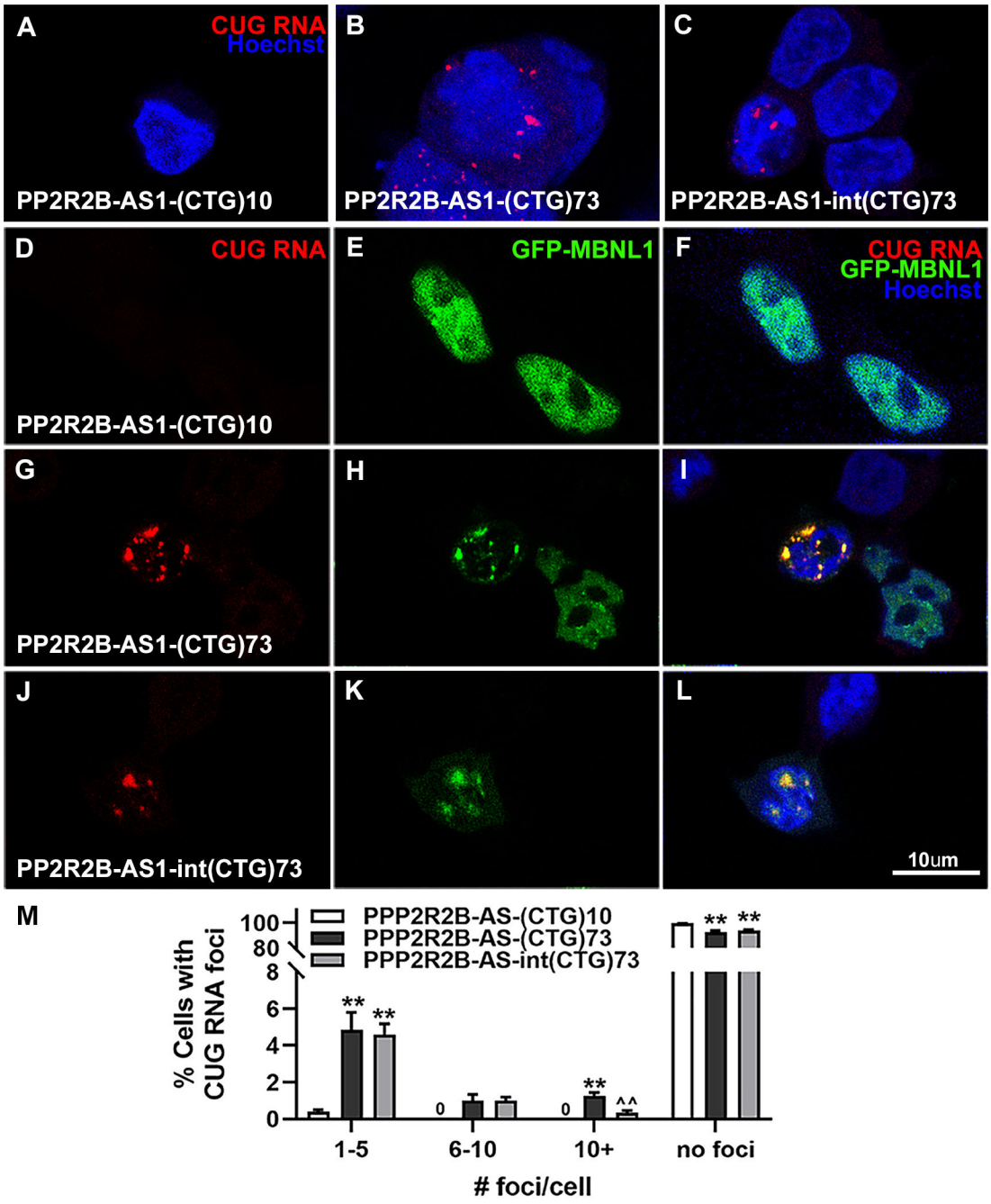

Fig. 3.

Expanded *PPP2R2B-AS1* transcripts form CUG RNA foci. (A-C) In situ hybridization using a CAG DNA/LNA probe indicates *PPP2R2B-AS1-(CUG)73* and *–int(CUG)73* transcripts form nuclear CUG RNA foci when transfected into SK-N-MC neuroblastoma cells. % of cells with foci and number of foci per cell were quantified (M). Mean ± SEM are shown. N=3 biological replicates, One-way ANOVA, *p<0.05, **p<0.01, compared with *PPP2R2B-AS1-(CTG)10*, ^^p<0.01, compared with *PPP2R2B-AS1-(CTG)73*. (D-L) GFP-MBNL1 shows nuclear diffused localization in cells that overexpress *PPP2R2B-AS1-(CUG)10* (D-F),

but becomes sequestered to CUG RNA foci formed by *PPP2R2B-AS1-(CUG)73* (G-I) or *-Int(CUG)73* (J-L) transcripts. Scale bar = 10um.

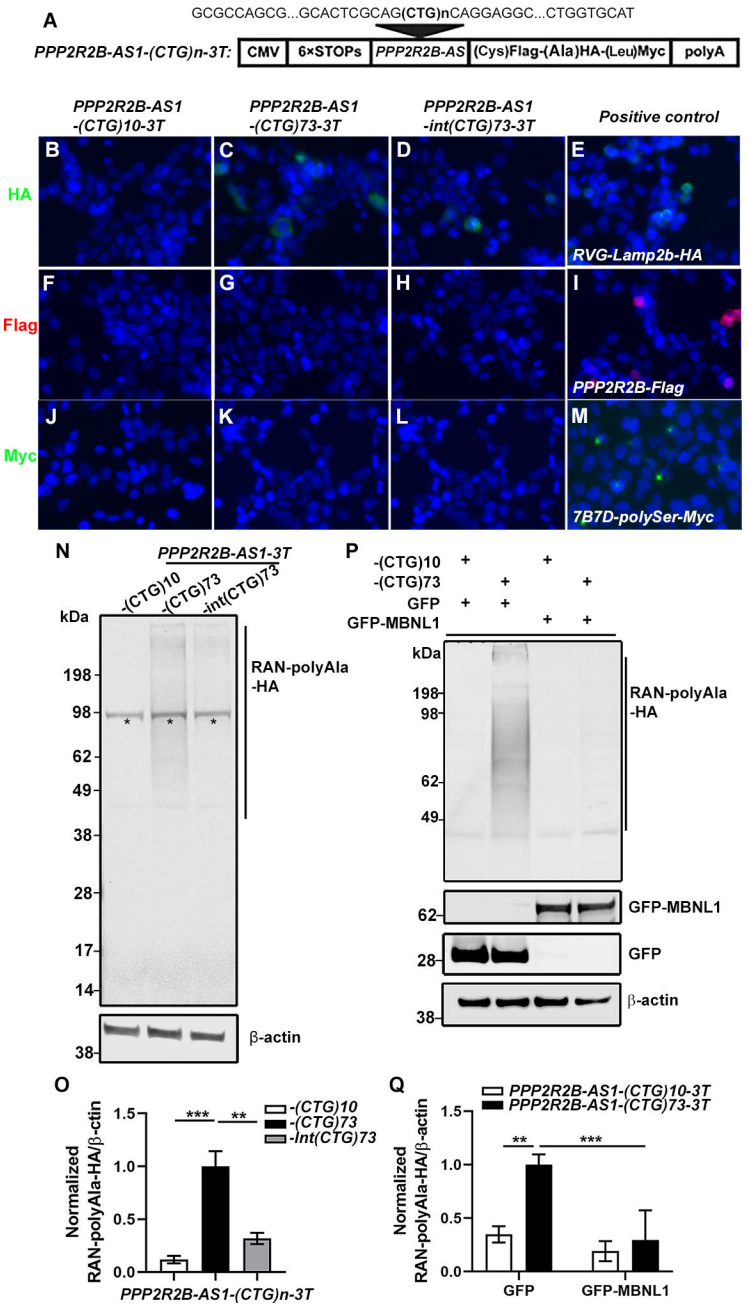

Fig. 4.

The expanded CUG repeat in *PPP2R2B-AS1* transcript may be translated into polyalanine (polyAla) containing proteins via RAN translation. (A) Schematic presentation of the tagged *PPP2RAB-AS*-(CTG)n-3T construct to determine the presence of RAN product. (B-I) Anti-HA staining revealed polyAla RAN protein in *PPP2R2B-AS1-(CTG)73–3T* and *PPP2R2B-AS1-int(CTG)73–3T* overexpressing SK-N-MC cells (C and D), compared with the *PPP2R2B-AS1-(CTG)10–3T* control cells (B), while anti-Flag or anti-Myc staining showed negative results (F-4 and J-4). SK-N-MC cells transfected with *RVG-Lamp2b-HA* (E), *PPP2R2B-Flag* (I), and *7B7D-polySer-Myc* (M), were used as positive controls for immunostaining using anti-HA, Flag, and Myc tag antibodies, respectively.

(N-O) RAN-polyAla-HA protein expression is decreased in *PPP2R2B-AS1-int(CTG)73–3T* overexpressing cells, compared with *PPP2R2B-AS1-(CTG)73–3T*. Asterisks mark nonspecific bands. N=4 biological replicates. The RAN-polyAla-HA/β-actin ratio in *PPP2R2B-AS1-(CTG)73–3T* transfected SK-N-MC cells was normalized to 1. Mean ± SEM are shown. (P-Q) GFP-MBNL1 overexpression decreased RAN-polyAla-HA expression in *PPP2R2B-AS1-(CTG)73* expressing SK-N-MC cells. N=3 biological replicates. The RAN-polyAla-HA/β-actin ratio in SK-N-MC cells co-transfected with *PPP2R2B-AS1-(CTG)73–3T* and *GFP* was normalized to 1. Mean ± SEM are shown. One-way ANOVA and Tukey posthoc test, *p<0.05, **p<0.01, ***p<0.001.

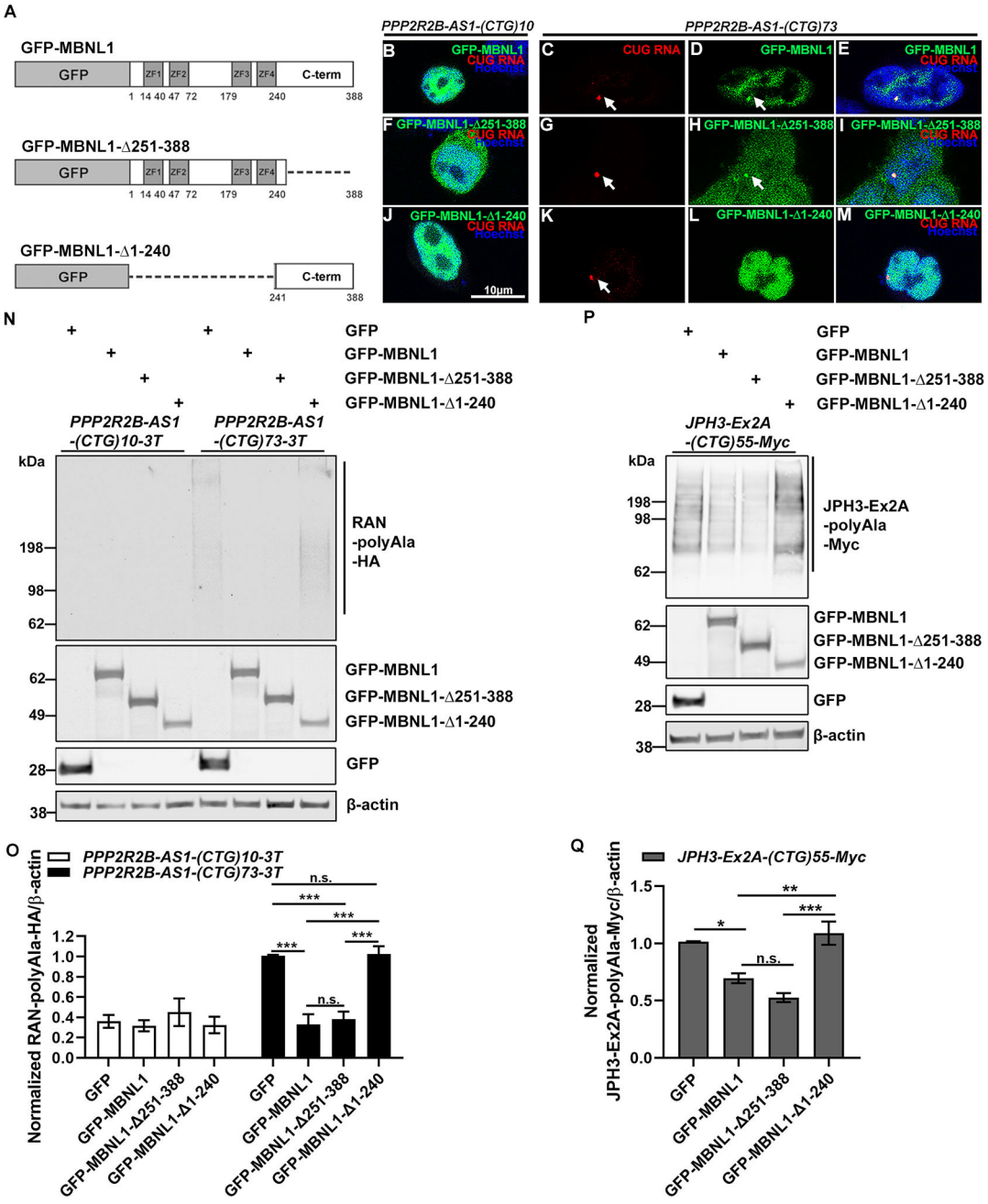

Fig. 5.

Effect of MBNL1 ZNFs on the production of RAN polyAla proteins from the *expPPP2R2B-AS1* transcript. (A). Schematic presentation of the GFP-MBNL1, GFP-MBNL1-Δ251–388, and GFP-MBNL1-Δ1–240 constructs. (B-M). *In situ* hybridization using a CAG DNA/LNA probe indicates exogenous GFP-MBNL1 and GFP-MBNL1-Δ251–388, but not GFP-MBNL1-Δ1–240, colocalizes with nuclear CUG RNA foci formed by the *PPP2R2R-AS1-(CUG)73* transcripts in SK-N-MC neuroblastoma cells. N=3 biological replicates. Representative images are shown. Scale bar = 10um. Arrows point to CUG RNA foci. (N-O) Overexpression of GFP-MBNL1 or GFP-MBNL1-Δ251–388, but not that of GFP-MBNL1-

Δ1–240, decreased RAN-polyAla-HA protein expression in *PPP2R2B-AS1-(CTG)73–3T* expressing SK-N-MC cells. N=4 biological replicates. The RAN-polyAla-HA/β-actin ratio in SK-N-MC cells co-transfected with *PPP2R2B-AS1-(CTG)73–3T* and *GFP* was normalized to 1. Mean ± SEM are shown. One-way ANOVA and Tukey posthoc test, *p<0.05, **p<0.01, ***p<0.001. (P-Q) Overexpression of GFP-MBNL1, or GFP-MBNL1-Δ251–388, but not that of GFP-MBNL1-Δ1–240, decreased polyAla-Myc expression in *JPH3-Ex2A-(CTG)55-Myc* expressing SK-N-MC cells. N=4 biological replicates, Mean ± SEM are shown. The JPH3-Ex2A-polyAla-Myc/β-actin ratio in SK-N-MC cells co-transfected with *JPH3-Ex2A-(CTG)55-Myc* and *GFP* was normalized to 1. One-way ANOVA and Tukey posthoc test, *p<0.05, **p<0.01, ***p<0.001.